CENTRE DE RECHERCHE
SUR LES TRANSPORTS
(C.R.T.)
CENTRE FOR RESEARCH
ON TRANSPORTATION

UNIVERSITÉ DE MONTRÉAL
ÉCOLE DES HAUTES ÉTUDES COMMERCIALES
ÉCOLE POLYTECHNIQUE

CENTRE DE RECHERCHE SUR LES TRANSPORTS
UNIVERSITÉ DE MONTRÉAL
C.P. 6128, SUCCURSALE A,
MONTRÉAL
CANADA H3C 3J7

TÉLÉPHONE: (514) 343-7575
TÉLEX: (21) 05-561325
EMAIL: crt@crt.umontreal.ca
TÉLÉCOPIEUR (FAX): (514) 343-7121

91-05261/

CRT-778A

Spatial Autoregressive Error Components in Travel Flow Models

by

Denis Bolduc[1]
Richard Laferrière[2]
Gino Santarossa[1]

[1] Département d'économique, Université Laval, Québec, Canada G1K 7P4

[2] Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale A, Montréal, Québec, Canada H3C 3J7

Centre de recherche sur les transports — Publication #778A
May 1991

Dépôt légal – 2ème trimestre 1991

RÉSUMÉ

Dans cette étude, nous proposons une généralisation de la formulation à composantes d'erreurs qui permet de capter les différents effets occasionnant la présence de corrélation dans les erreurs de modèles de régression avec données de flux. Selon la formulation proposée, le terme d'erreur se décompose en une somme d'une erreur relative à la zone d'origine, une erreur relative à la zone de destination et une erreur associée au flux. Chaque composante d'erreur est issue d'un processus générateur autorégressif spatial d'ordre 1. L'estimation des paramètres du modèle est basée sur la méthode du maximum de vraisemblance. La méthodologie proposée a l'avantage de demeurer applicable même dans le contexte d'échantillons de grande taille.

Mots-clés: modèle d'erreurs composées généralisées, moindres carrés généralisés, autocorrélation spatiale, modèle d'explication de flux

ABSTRACT

In this study, we propose a generalization of the error component formulation to model the correlation among the errors of a regression based on travel flow data. The error term is broken down into a sum of one component related to the origin zone, one component related to the destination zone and a remainder. Each component is assumed to result from a first-order spatial autoregressive generating process. An efficient estimation approach based on maximum likelihood which addresses the practical implementation of such a model with a large sample size is suggested.

Key-words: Generalized Error Components, Generalized Least-Squares, Maximum Likelihood, Spatial Autocorrelation, Travel flow Modeling.

Table of Contents

1 INTRODUCTION .....

2 TRAVEL FLOW MODELS .....

3 THE ERROR FORMULATION WITH FIRST-ORDER SPATIALLY AUTO-REGRESSIVE COMPONENTS 2
3.1 Some Sub-Models 5
- The Brandsma and Ketellapper formulation 5
- The Bolduc, Dagenais and Gaudry formulation 5

4 ESTIMATION ISSUES 6
4.1 An Efficient Maximum Likelihood Algorithm for Large Problems 8

5 CONCLUSION 9

REFERENCES 10

1 INTRODUCTION

In this paper we introduce a flexible framework to deal with the problem of spatial autocorrelation among the errors of a regression based on travel flow data. This new model formulation integrates first-order spatial autoregressive processes with standard error components to allow one to account for the several sources of misspecification generally associated with this type of models in a natural way. The error term is broken down into a sum of one component related to the zone in origin, one component associated with the zone in destination and a remainder. Each component is assumed to result from a first-order spatial autoregressive generating process. We refer to this model formulation as a travel flow regression model with first-order spatially autoregressive error components (EC-SAR(1)).

Brandsma and Ketellapper (BK (1)) and Bolduc, Dagenais and Gaudry (BDG (2)) have proposed solutions to this problem. Both approaches rely on spatially autoregressive errors. BK's solution ignores a very important constituent of the error term. BDG extends the latter formulation to also incorporate this missing effect into the model. Their proposed strategy is to model the inter-dependencies between the flow errors using a conventional first-order spatial autoregressive error process with a weighting matrix properly defined. However, except for very simple weight specifications, the correlation effects captured by the error formulation may sometimes be complicated to interpret. EC-SAR(1) differs from these methodologies in offering a simpler and more natural framework to handle the spatial autocorrelation among the errors. The suggested approach includes the two previous formulations as special cases. The key improvement with the EC-SAR(1) approach is in the way the specifics of flows data are accounted for.

In the next section, we summarize the main concepts used in the study field of flow data. Section 3 introduces the travel flow model with SAR(1) error components. We use a notation which can handle both cases of balanced and unbalanced data. Then, the methodologies suggested in BK and BDG are shown to be sub-models of the more general framework proposed here. In the last section, we discuss a critical issue regarding the practical estimation of this model with a large sample size and propose an efficient maximum likelihood algorithm as a solution. The last section concludes.

2 TRAVEL FLOW MODELS

A regression model based on travel flow data may be written as:

$$y_{ij} = \beta_0 + N_{ij}\beta_1 + S_i\beta_2 + S_j\beta_3 + \varepsilon_{ij}, \quad i = 1,\ldots,N, \quad j = 1,\ldots,T, \tag{1}$$

where y_{ij} is a dependent variable which represents a flow between a given pair of zones i and j. The index i refers to the zones belonging to the set of N zones in origin and the index j represents one of the T zones at the destination. The same notation is used to model both *inter* urban flow data where the originating zones are distinct form the zones at the destination and

intra urban (interzonal) flow data where the set of zones in origin is identical to the set of zones in destination. In that particular case, N would be equal to T. In the following, we keep the notation general by allowing different sets of zones in origin and in destination.

Passengers' travel flow regression models typically contain three kinds of variables: network variables (N_{ij}), socio-economic variables associated with the origin (S_i) and socio-economic variables associated with the destination (S_j). Network variables usually refer to levels of service (travel time, fare, frequency, etc.) of the relevant modes and socio-economic variables (population, income, employment, etc.) characterize the generation and attraction capacity of the two zones involved in the flow. β_0 is a constant term and β_1, β_2 and β_3, are vectors of fixed coefficients associated with the variables described above.

Note that a given observation of the N_{ij}, S_i and S_j variables refers only to the corresponding *ij* market, that is network variables for the competing markets *is* (same origin but alternative destination) or *rj* (same destination but alternative origin) do not appear in (1). The same is true for socio-economic variables of zones other than i and j. It might be conceivable that socio-economic and network variables adjacent to market *ij* be incorporated in the relationship explaining y_{ij}, but this could well lead to an intractable model. The omission of those variables gives raise to some spatial autocorrelation in the regression errors. Sources of spatial autocorrelation among errors are model misspecification and omission of explanatory variables capturing effects related to the physical and economic structure (distance between zones, size of zones, length of frontier between adjacent zones, etc.) of the region. For these reasons, the error term ε_{ij} in equation (1) will reflect this situation. Cliff and Ord (3) suggest that this inter-relationship is caused by the dependent nature of the zones belonging to a region.

We argue that a suitable formulation to capture spatial autocorrelation among this type of residual error should incorporate three elements: one related to the origin zones, one to the destination zones and the other to the network specificities. The model formulation that we suggest in this paper captures these different effects in an appropriate way. This specification is formalized in the next section.

3 THE ERROR FORMULATION WITH FIRST-ORDER SPATIALLY AUTOREGRESSIVE COMPONENTS

We propose to define an error structure that takes into account the several sources of misspecification discussed above. The formulation is a generalization of the error component decomposition where the ε_{ij} error term in equation (1) is written as

$$\varepsilon_{ij} = \alpha_i + \lambda_j + u_{ij}. \qquad (2)$$

This decomposition of the unexplained flow between zones i and j is consistent with the specification of the fixed part of the regression model. The error term α_i is interpreted as capturing unmeasured effects associated with origin i (it contains omitted variables from the S_i vector), λ_j picks up those missing effects related to destination j (it contains omitted variables from the S_j vector) and finally u_{ij} incorporates all the network variables not included into the model specification. The component α_i is a random term assumed to be of zero mean. To allow for the presence of spatial correlation in origin, α_i is assumed to be correlated with the error α_r of an alternative zone in origin. Similar assumptions are made for the two other error components. Also, we assume that all these terms are uncorrelated with each other. This is written as:

$$E(\alpha_i\lambda_j) = E(\alpha_i u_{ij}) = E(\lambda_j u_{ij}) = 0 \quad \forall i,j. \tag{3}$$

To summarize the several sources of spatial autocorrelation, we postulate a normalized first-order generalized spatial autoregressive (SAR(1)) generating process for each component. We write it as

$$\alpha_i = \rho_1 \sum_{r=1}^{N} w_{1,ir}\alpha_r + \xi_i, \tag{4}$$

$$\lambda_j = \rho_2 \sum_{s=1}^{T} w_{2,js}\lambda_s + \zeta_j, \tag{5}$$

$$u_t = \rho_3 \sum_{l=1}^{L} w_{3,tl}u_l + \upsilon_t, \tag{6}$$

where for notational simplicity, t designates a given flow $(i \rightarrow j)$ and l an alternative flow $(r \rightarrow s)$. We postulate that $\xi_i \sim N(0,\sigma_\xi^2)$, $\zeta_j \sim N(0,\sigma_\zeta^2)$ and $\upsilon_t \sim N(0,\sigma_\upsilon^2)$ $\forall i, \forall j$ and $\forall t$, respectively. These formulations translate the interdependencies discussed above into an operational framework. The processes are normalized so that $\sum_r w_{1,ir} = 1$, $\forall i$, $\sum_s w_{2,js} = 1$, $\forall j$ and $\sum_l w_{3,tl} = 1$, $\forall t$. The parameter ρ_1 is a spatial autocorrelation coefficient which measures the degree of linear dependency between α_i and each other "neighbor" α_r in origin. The parameters ρ_2 and ρ_3 are defined analogously. These coefficients are restricted to lie within the unit interval $(-1,1)$ in order to ensure the stability of each of the processes. The stability of the SAR(1) processes is required to guarantee that the objective function in the maximum likelihood framework is continuous in the model parameters over the domain defined by the three correlation coefficients. Recall that continuity in parameters of the optimizing function is critical to produce consistent and asymptotically normally distributed estimators. An alternative normalization can define a different stability range for ρ. Since the unit interval is very intuitive, we will restrict our analysis to the chosen normalization.

To start with, we describe the weights used in the autoregressive process associated with the origin. The weights $w_{1,ir}$ are ***estimable*** functions specified by the analyst to describe the influence of each origin specific error upon the others. Using a general notation, we write this function as

$$w_{1,ir} = \frac{w^*_{1,ir}}{\sum_{r=1}^{N} w^*_{1,ir}}, \quad w_{1,ii} = 0 \quad \forall i, \tag{7}$$

$$w^*_{1,ir} = g\left(R^1_{1,ir}, R^2_{1,ir}, \ldots, R^{L_1}_{1,ir}, \theta_{11}, \theta_{12}, \ldots, \theta_{1H_1}\right), \quad w^*_{1,ii} = 0 \quad \forall i, \tag{8}$$

where there are H_1 parameters θ_{1h} and L_1 variables $R^l_{1,ir}$. The weight $w_{1,ir}$ measures the importance of α_r in explaining α_i. The larger $w_{1,ir}$ is, the more are α_i and α_r related. The process is said to be generalized because the θ_{1h} parameters are estimated. A current practice in spatial studies is to fix the value of these parameters prior to estimation. One simple weight definition is $w^*_{1,ir} = (d_{ir})^{-\alpha}$ where $\alpha > 0$ and where d_{ir} is the distance between zones i and r. Applied to our model, this formulation is one of the first candidate which can be used to capture the correlation between the zone specific errors. The interpretation and the sign of ρ_1 depend on the definition of proximity embodied in $w^*_{1,ir}$. For the specification just discussed, a $\rho_1 > 0$ implies that errors of the same sign are geographically grouped together. Alternatively, had we specified $w^*_{1,ir} = (P_i P_r)^{\alpha}$, $\alpha > 0$, where P_i and P_r are populations (or incomes) in zone i and r respectively, we would have found the strongest correlations among the errors of zones containing the largest populations (incomes). Spatial studies that use Spatial AutoRegressive of order 1 (SAR(1)) error processes often define the contiguity structure through a Boolean contiguity matrix. In that case, one would have $w^*_{1,ir} = 1$ if i and r are contiguous and $w^*_{1,ir} = 0$ otherwise. (See Anselin (4), Bivand (5), Blommestein (6) or Cliff and Ord (3), for more details. For a recent application of SAR(1) processes in economics, see Case (7) and in discrete choice analysis, see Bolduc and Ben-Akiva (8)).

An identical discussion could also be made for the SAR(1) process associated with the destination. To simplify, we skip to the third error component. The interpretation of the SAR(1) process pertaining to the network specific effects is slightly different. Several functional forms for $w_{3,il}$ were considered in BDG. To shorten the discussion, the reader is referred to that paper for details. A particular specification will be described in the next section where we show that the alternative error formulations used in BK and BDG are sub-models of our proposed error generating process. They are obtained with $\sigma^2_\xi = 0$ and $\sigma^2_\zeta = 0$. This essentially eliminates the first two error components. Another interesting sub-model is the standard two-way error component formulation which is obtained with $\rho_1 = \rho_2 = \rho_3 = 0$.

3.1 Some Sub-Models

- The Brandsma and Ketellapper formulation

A solution to the problem of spatial correlation among flow errors is the spatial autoregressive process of order two (SAR(2)) proposed by Brandsma and Ketellapper (1), which can be written as

$$u_t = \gamma_1 \sum_{l=1}^{L} w_{tl}^{o} u_l + \gamma_2 \sum_{l=1}^{L} w_{tl}^{d} u_l + \upsilon_t, \quad \upsilon_t \sim N(0, \sigma_\upsilon^2), \tag{9}$$

with weights normalized according to equation (7) where

$w_{tl}^{o*} = 1$ if flows t and l share a same origin,

$= 0$ otherwise, and

$w_{tl}^{d*} = 1$ if flows t and l share a same destination,

$= 0$ otherwise.

To see that this formulation is a special case of our proposed framework, rewrite this last generating process into a first-order autoregressive process as follows

$$u_t = \gamma_1 \sum_{l=1}^{L} \left(w_{tl}^{o} + \frac{\gamma_2}{\gamma_1} w_{tl}^{d} \right) u_l + \upsilon_t, \quad \upsilon_t \sim N(0, \sigma_\upsilon^2). \tag{10}$$

It can be interpreted as an EC-SAR(1) generating process with no origin and no destination specific error component. It defines a very peculiar weight specification for the third component of the process where only origin and destination effects are accounted for. The formulation can be criticized in that it incorporates the zone specific error correlation effects at the wrong place.

- The Bolduc, Dagenais and Gaudry formulation

This formulation extends the last one by including extra terms to capture network specific factors missing from the fixed part of the model. The most simple spatial correlation structure that they suggest has an autoregressive process of order 1 (SAR(1)) structure

$$u_i = \rho \sum_{l=1}^{L} w_{il}\, u_l + \upsilon_i, \quad \upsilon_i \sim N(0, \sigma_\upsilon^2), \tag{11}$$

where the non normalized weight is defined as

$$w_{il}^{*} = (d_{ir} + d_{js})^{-\theta_1} + (d_{is} + d_{rj})^{-\theta_2}, \quad \text{with} \quad w_{il} = 0 \quad \forall i. \tag{12}$$

This function takes into consideration direct and cross distance effects between two origins and two destinations. For a given pair of travel flows say $(i \rightarrow j)$ and $(r \rightarrow s)$, the first of the two components in equation (12), the direct effect, refers to the distances between the two origins i and r and the two destinations j and s. For a given value of $\theta_1 > 0$, small distances between the origins and/or the destinations induce a large weight w_{il}. This first term is used to account for a mixture of positive and negative relationships between individual flows originating and terminating in specified zones. The second term in the RHS of equation (12) represents the cross distance effect between the origins and the destinations. The formulation implies that for a given $\theta_2 > 0$, everything else being equal, the higher those distances are, the smaller is the weight w_{il}. A more detailed justification for this formulation may be found in BDG.

This generating process is clearly a sub-case of our proposed specification. Although it incorporates the three major components for explaining spatial autocorrelation, these correlation effects are not included in the most appropriate way. For more complicated specifications, the effects captured by the weights may sometimes be complicated to interpret. Our modeling strategy is to put in the third component of the EC-SAR(1) formulation only those effects that do not belong to the zone specific components. This way the interpretation of the weight in the last error component will be simplified.

4 ESTIMATION ISSUES

In order to discuss estimation issues, it is more appropriate to rewrite the EC-SAR(1) formulation in matrix notation. Incorporating all network specific and socio-economic variables into a single matrix denoted by X, the regression model with EC-SAR(1) errors is written as:

$$y = X\beta + \varepsilon, \tag{13}$$

$$\varepsilon = C\alpha + D\lambda + u, \tag{14}$$

with

$$\alpha = \rho_1 W_1 \alpha + \xi, \quad \xi \sim N(0, \sigma_\xi^2 I_N), \tag{15}$$

$$\lambda = \rho_2 W_2 \lambda + \zeta, \quad \zeta \sim N(0, \sigma_\zeta^2 I_T), \tag{16}$$

$$u = \rho_3 W_3 u + \upsilon, \quad \upsilon \sim N(0, \sigma_\upsilon^2 I_L), \tag{17}$$

where L denotes the total number of observations in the sample. The vectors y, ε, u and υ are of dimension $L \times 1$. The vectors α and ξ are N-dimensional and λ and ζ are T-dimensional. X is a $L \times K$ matrix of fixed regressors and β is a $K \times 1$ vector of parameters. The weights defined in equations (4)-(6) are incorporated into the matrices W_1, W_2 and W_3 of dimension $N \times N$, $T \times T$ and $L \times L$, respectively. These matrices are normalized so that the corresponding SAR(1) processes are stable over the unit $(-1, 1)$ interval of the ρ coefficient associated with the process.

With balanced data, (i.e. each origin has the same set of destinations and each destination has the same set of origins), $C = (I_N \otimes e_T)$ and $D = (e_N \otimes I_T)$ where e_T denotes a T dimensional unit vector and I_T an identity matrix of size T. With unbalanced data, the C and D matrices need to be redefined in order to account for the missing flow observations. Define T_i as the number of destinations associated with origin i. T_i can be computed as $T - z_i$ where z_i is the number of missing destinations from origin i. Denote e_{T_i} as a $T_i \times 1$ unit vector and $I_{(T_i)}$ as an identity matrix of size $T \times T$ with rows associated with missing destinations from origin i eliminated. Then one can easily check that C and D with unbalanced data can be written as

$$C = \begin{bmatrix} e_{T_1} & 0 & \dots & 0 \\ 0 & e_{T_2} & \dots & 0 \\ . & . & \dots & . \\ . & . & \dots & . \\ 0 & 0 & \dots & e_{T_N} \end{bmatrix}_{L \times N} \quad \text{and} \quad D = \begin{bmatrix} I_{(T_1)} \\ I_{(T_2)} \\ . \\ . \\ I_{(T_N)} \end{bmatrix}_{L \times T}.$$

No other modification need to be made to the formulation to account for the missing observations.

Based on the assumptions made for the autoregressive processes, equations (15)-(17) may be written as:

$$\alpha = (I_N - \rho_1 W_1)^{-1} \xi = P_1^{-1} \xi, \tag{18}$$

$$\lambda = (I_T - \rho_2 W_2)^{-1} \zeta = P_2^{-1} \zeta, \tag{19}$$

$$u = (I_L - \rho_3 W_3)^{-1} \upsilon = P_3^{-1} \upsilon, \tag{20}$$

where we use $P_1 = I_N - \rho_1 W_1$, $P_2 = I_T - \rho_2 W_2$ and $P_3 = I_L - \rho_3 W_3$.

Substituting equations (18)-(20) into equation (14), the complete model may be written as:

$$y = X\beta + \varepsilon \tag{21}$$

$$\varepsilon = CP_1^{-1}\xi + DP_2^{-1}\zeta + P_3^{-1}\upsilon. \tag{22}$$

Our distributional assumptions imply that $\varepsilon \sim N(0, \Phi)$, with

$$\Phi = \sigma_\xi^2 CP_1^{-1}P_1^{-1\prime}C' + \sigma_\zeta^2 DP_2^{-1}P_2^{-1\prime}D' + \sigma_\upsilon^2 P_3^{-1}P_3^{-1\prime},$$

a $L \times L$ covariance matrix. To estimate the model, we use the maximum likelihood framework. In the following discussion we assume that the data come from a balanced data setting. For a fairly small problem with 20 origins and 20 destinations, for example, Φ and P_3 would be of dimension 400×400. Note that in the implementation of the maximum likelihood algorithm, those matrices need to be inverted at each iteration. Although 400×400 is still computationnaly tractable, for larger problems, the solution may become numerically unreachable in a reasonable amount of time. In the following, we rewrite the model formulation in an equivalent form where only one matrix of size $L \times L$ needs to be inverted. This inverse will be based on the inverse of a $(N+T) \times (N+T)$ matrix. This may represent a huge computational gain in larger problems.

4.1 An Efficient Maximum Likelihood Algorithm for Large Problems

First, to simplify the notation we rewrite the model as:

$$y = X\beta + \varepsilon, \tag{23}$$

$$\varepsilon = Ma + P_3^{-1}\upsilon, \quad \upsilon \sim N(0, \sigma_\upsilon^2 I_L), \tag{24}$$

where $M = [\sigma_\xi CP_1^{-1} \mid \sigma_\zeta DP_2^{-1}]$ is a matrix of dimension $L \times (N+T)$. The operator $\mid$ designates the horizontal concatenation of two matrices. The vector a is $(N+T)$-dimensional with elements distributed as independent standard normal variates. We write it as $a \sim N(0, I_{N+T})$. To simplify the numerical calculations involved in producing the maximum likelihood estimators, we rewrite the error process into an equivalent form. Define $M^* = P_3 M$. This yields

$$\varepsilon = P_3^{-1}[P_3 Ma + \upsilon] = P_3^{-1}[M^* a + \upsilon] = P_3^{-1}\omega, \tag{25}$$

where $\omega \sim N(0, \Omega)$ with $\Omega = [M^* M^{*\prime} + \sigma_\upsilon^2 I_L]$. Therefore, $\text{cov}(\varepsilon) = \Phi = P_3^{-1}\Omega P_3^{-1\prime}$, with an inverse calculated as $\Phi^{-1} = P_3'\Omega^{-1}P_3$, a $L \times L$ matrix. To compute Ω^{-1}, we use some well-known formulas which in this case make it depend only on the inverse of a $(N+T) \times (N+T)$ matrix. It yields

$$\Omega^{-1} = \sigma_v^{-2}[I_L - M^* V^{-1} M^{*\prime}], \tag{26}$$

where $V = [\sigma_v^2 I_{N+T} + M^{*\prime} M^*]$, is a $(N+T)$-dimensional square matrix. The computation of the determinant of Ω can also be simplified using the result

$$|\Omega| = (\sigma_v^2)^{L-(N+T)} |V|. \tag{27}$$

With the above formulas and denoting γ as the joint vector of parameters involved in the computation of Φ, the likelihood function may be written as:

$$l(\beta, \gamma | y) = (2\pi)^{-L/2} |P_3| |\Omega|^{-1/2} \exp\left\{-\frac{(y - X\beta)' P_3' \Omega^{-1} P_3 (y - X\beta)}{2}\right\},$$

$$= (2\pi)^{-L/2} |P_3| |\Omega|^{-1/2} \exp\left\{-\frac{\omega' \Omega^{-1} \omega}{2}\right\}. \tag{28}$$

The corresponding log-likelihood function is finally written as:

$$L(\beta, \gamma | y) = -\frac{L}{2} \ln(2\pi) + \ln |P_3| - \frac{1}{2} \ln |\Omega| - \frac{\omega' \Omega^{-1} \omega}{2}. \tag{29}$$

The reader may verify that the optimization function (29) can be concentrated with respect to β. This brings the problem down to a maximization of (29) with respect to the vector of nuisance parameters γ, only.

5 CONCLUSION

In order to approximate the spatial correlation among the errors of a regression model on flow data, we have proposed a new error generating process based on the theory of error components and on spatial autoregressive of order 1 processes. The error term is broken down into a sum of one component related to the origin zone, one component related to the destination zone and a remainder. Each component is assumed to result from a first-order spatial autoregressive generating process. An efficient estimation approach based on maximum likelihood which addresses the practical implementation of such a model with a large sample size has been suggested.

REFERENCES

1.Brandsma, A.S. and R.H. Ketellapper, 1979,"A Biparametric Approach to Spatial Autocorrelation", *Environment and Planning*, A, 11, 51-58.

2.Bolduc, D., M. G. Dagenais and M.J. Gaudry, 1989,"Spatially Autocorrelated Errors in Origin-Destination Models: A New Specification Applied to Urban Travel Demand in Winnipeg", *Transportation Research*, B, 23, 5, 361-372.

3.Cliff A.D. and J.K. Ord, 1981, *Spatial Processes, Models and Application*, Pion, London.

4.Anselin, L., 1989, *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers.

5.Bivand, R.S., 1984, "Regression Modeling with Spatial Dependence: an Application of Some Class Selection and Estimation Methods", *Geographical Analysis*, 16, 1(January), 25-37.

6.Blommestein, H.J., 1983, "Specification and Estimation of Spatial Econometric Models. A Discussion of Alternative Strategies for Spatial Economic Modeling", *Regional Science and Urban Economics*, 13, 251-270.

7.Case, A., 1991, "Spatial Correlation in Household Demand", forthcoming in Econometrica.

8.Bolduc D. and Ben-Akiva M., 1991, "A Multinomial Probit Formulation for Large Choice Sets", *Proceedings of the 6th International Conference on Travel Behavior*, May 22, 23, 24, 1991, Quebec.

SPATIAL AUTOREGRESSIVE ERROR COMPONENTS IN TRAVEL FLOW MODELS

by

BOLDUC Denis
Département d'économique,
Université Laval, Québec, G1K 7P4, Canada.

LAFERIERRE Richard
Centre de recherche sur les transports,
Université de Montréal, C.P. 6128 Succ. A, Montréal, Canada.

and

SANTAROSSA Gino
Département d'économique,
Université Laval, Québec, G1K 7P4, Canada.